Environmental Health Perspectives
Vol. 96, pp. 101–111, 1991

A Proposed Method for Assembly and Interpretation of Short-Term Test Data

by David Brusick*

The genetic toxicology databases for chemicals that have been tested extensively are generally composed of inconsistent responses from a diverse set of assays. Consequently, difficulties arise when the data are evaluated for classifying the agent or for assessing the chemical's hazard potential. Several years ago, the International Commission for Protection against Environmental Mutagens and Carcinogens (ICPEMC) established a committee to construct a process for compiling and interpreting diverse data sets. The Committee has developed a weight-of-evidence approach that combines test data into a series of scores for test type, class, family, and a consensus score defining the relative mutagenic activity of the agent compared with other chemicals in the database. This report describes the method and preliminary results from 113 chemicals.

Introduction

Committee 1 of The International Commission for Protection against Environmental Mutagens and Carcinogens (ICPEMC) was established in 1979 to review the status of short-term tests for mutagenicity and the degree to which these tests are concordant with results from three mammalian *in vivo* tests (dominant lethal, heritable translocation, and specific locus) measuring germ cell damage (*1*). The mission of committee 1 was broadened in 1983 to develop, if possible, a method that would integrate and interpret results from heterogeneous data typical of mutagenicity test batteries.

Committee 1 members began with a weight-of-evidence scheme proposed by Brusick (*2*). This system was based on a method of weighted averages of both positive and negative test results from a battery consisting of both *in vitro* and submammalian assays. Allthough the committee retained the weight-of-evidence portion of the approach, range of assays and the mechanics of data handling for the current method have evolved substantially.

There were three primary objectives that committee 1 set out to accomplish in the design of a data analysis method. The first goal was to develop a method that would extend the use of a database beyond listing tests and results. For mutagenicity there was a need for a process to assemble the test results for a chemical in a manner that would produce a consensus regarding the mutagenic activity of the agent. The second goal was to use the results of the evaluations to rank chemicals and compare that rank order with other properties of the same chemicals such as cancer or germ cell mutation. The third goal was to use the data analysis with a large database to understand mutagenicity tests and their relationships to each other and to chemicals and chemical classes.

*Hazleton Washington, Inc., 9200 Leesburg Pike, Vienna, VA 22182.

Comparison with Other Methods of Data Analysis

Several other investigators have developed or proposed approaches to accomplish many of the objectives stated above. One of the earlier uses of the data in this manner was proposed by Squire (*3*) in which he suggested a semiquantitative approach that estimated carcinogenic potential using a point system for various characteristics of a chemical. Mutagenicity was highest weighted of all components of his carcinogen prediction scheme.

In the mid-1980s, Waters et al. (*4*) developed a linear profile of mutagenic activity that illustrated the positive and negative results for all tests conducted on a chemical (Fig. 1). This plot, identified as a Genetic Activity Profile (GAP), has undergone several improvements and is currently available with an extensive database on PC-based software (*4*). GAPs facilitate direct comparison of test responses for chemicals of similar classes and/or structural relatedness.

Other investigators have attempted to used statistical (*5*) and structure–activity analyses (*6*) of short-term test results to predict carcinogenic activity of chemicals and to construct more reliable

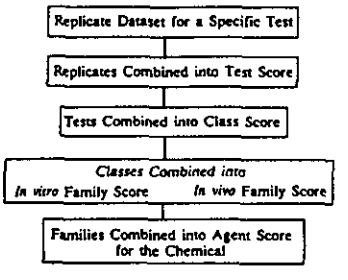


FIGURE 1. The general process of data reduction from individual trails to a single agent score. The merging process at each step involves simple averaging of scores.

Table 1. The current family, class, and test structure.[a]

Family: *In vitro*
Class A1: Primary DNA damage—prokaryotes
- BRD Other DNA repair-deficient bacteria differential toxicity
- BSD *Bacillus subtilis rec* strains, differential toxicity
- ECD *Escherichia coli pol*A/W3110-P3478, differential toxicity (spot test)
- ECL *Escherichia coli pol*A/W3110-P3478, differential toxicity (liquid suspension)
- ERD *Escherichia coli rec* strains, differential toxicity

Class A2: Primary DNA damage—lower eukaryotes
- SCG *Saccharomyces cerevisiae*, gene conversion
- SCH *S. cerevisiae*, homozygosis by recombination or gene conversion

Class A3: Primary DNA damage—mammalian cells
- UHF Unscheduled DNA synthesis, human fibroblasts
- UHL Unscheduled DNA synthesis, human lymphocytes
- UHT Unscheduled DNA synthesis, transformed human cells
- UIA Unscheduled DNA synthesis, other animal cells
- UIH Unscheduled DNA synthesis, other human cells
- URP Unscheduled DNA synthesis, rat primary hepatocytes

Class A4: Gene mutation—prokaryotes
- BSM *Bacillus subtilis* multigene test
- EC2 *Escherichia coli*WP2, reverse mutation
- ECF *Escherichia coli* (excluding strain K12), forward mutation
- ECK *Escherichia coli* WP2 UVRA, reverse mutation
- ECR *Escherichia coli* (other miscellaneous strains), reverse mutation
- SAL *Salmonella typhimurium*, all strains

Class A5: Gene mutation—lower eukaryotes
- NCF *Neurospora crassa*, forward mutation
- NCR *Neurospora crassa*, reverse mutation
- SCF *Saccharomyces cerevisiae*, forward mutation
- SCR *Saccharomyces cerevisiae*, reverse mutation
- SZF *Schizosaccharomyces pombe*, forward mutation

Class A6: Gene mutation—mammalian cells
- G51 Gene mutation, mouse L5178Y, all other loci
- G5T Gene mutation, mouse L5178Y cells, *TK* locus
- G9H Gene mutation, Chinese hamster lung V-79 cell, HPRT
- G9O Gene mutation, Chinese hamster lung V-79 cells, ouabain
- GCO Gene mutation, Chinese hamster ovary cells
- GIA Gene mutation, other animal cells

Class A7: Aneuploidy—lower eukaryotes
- SCN *Saccharomyces cerevisiae*, aneuploidy

Class A8: Sister chromatid exchange—mammalian cells
- SIA Sister chromatid exchange, other animal cells
- SHL Sister chromatid exchange, human lymphocytes
- SHF Sister chromatid exchange, human fibroblasts
- SIC Sister chromatid exchange, Chinese hamster cells
- SIH Sister chromatid exchange, other human cells
- SIS Sister chromatid exchange, Syrian hamster cells
- SIR Sister chromatid exchange, rat cells
- SIM Sister chromatid exchange, mouse cells
- SIT Sister chromatid exchange, transformed cells

Class A9: Chromosome aberration—mammalian cells
- CIA Chromosome aberrations, other animals cells
- CHL Chromosome aberrations, human lymphocytes
- CHF Chromosome aberrations, human fibroblasts
- CIC Chromosomal aberrations, Chinese hamster cells
- CIH Chromosomal aberrations, other human cells
- CIS Chromosomal aberrations, Syrian hamster cells
- CIR Chromosomal aberrations, rat cells
- CIT Chromosomal aberrations, transformed cells

Class A10: Transformation—mammalian cells
- T7S Cell transformation, SA7/Syrian hamster embryo cells
- T7R Cell transformation, SA7/rat cells
- TBM Cell transformation, BALB/C3T3 mouse cells
- TCL Cell transformation, other established cell lines
- TCS Cell transformation, Syrian hamster embryo cells, clonal assay
- TCM Cell transformation, C3H10T1/2 mouse cells
- TRR Cell transformation, RLV/Fischer rat embryo cells

Family: *In vivo*
Class B1: DNA repair, somatic—mammal
- UBH Unscheduled DNA synthesis, human bone marrow cells
- UPR Unscheduled DNA synthesis, rat hepatocytes
- UVA Unscheduled DNA synthesis, other animal cells
- UVC Unscheduled DNA synthesis, hamster cells
- UVR Unscheduled DNA synthesis, other rat cells
- UVM Unscheduled DNA synthesis, mouse cells

Class B2: Gene mutation, somatic—insect Drosophila
- DMM *Drosophila melanogaster*, somatic mutation (and recombination)

Class B3: Spot test, somatic—mammal
- MST Mouse spot test

Class B4: Sister chromatid exchange, somatic—mammal
- SLH Sister chromatid exchange, human lymphocytes
- SVA Sister chromatid exchange, animal cells
- SVH Sister chromatid exchange, other human cells

Class B5: Micronuclei, somatic—mammal
- MVC Micronucleus test, hamsters
- MVM Micronucleus test, mice
- MVR Micronucleus test, rats

Class B6: Chromosome aberration, somatic—mammal
- CBA Chromosomal aberrations, animal bone marrow cells
- CBH Chromosomal aberrations, human bone marrow cells
- CLA Chromosomal aberrations, animal leukocytes
- CLH Chromosomal aberrations, human lymphocytes
- CVA Chromosomal aberrations, other animal cells treated

Class B7: Heritable damage—insect, Drosphila
- DMH *Drosophila melanogaster*, heritable translocation test
- DML *Drosphila melanogaster*, dominant lethal test
- DMX *Drosophila melanogaster*, sex-linked recessive lethal mutation

Class B8: Heritable specific locus—mammal
- SLO Mouse specific locus test, other stages

Class B9: Dominant lethal—mammal
- DLM Dominant lethal test, mice
- DLR Dominant lethal test, rats

Class B10: Heritable translocation—mammal
- MHT Mouse heritable translocation test

Class B11: Chromosome aberration, germinal—mammal
- CCC Chromosomal aberrations, spermatocytes treated and observed
- CGC Chromosomal aberrations, spermatogonia treated and observed
- CGG Chromosomal aberrations, spermatogonia treated and observed
- COE Chromosomal aberrations, oocytes or embryos treated

Class B12: Sperm morphology—mammal
- SPM Sperm morphology, mouse
- SPR Sperm morphology, rat
- SMS Sperm morphology, sheep

[a]Code names according to Waters et al. (*4*). Only 85 tests are used. Critrion has been that in the total database a test had to be used for at least five chemicals.
[b]All strains of Salmonella included. The highest dose negative or lowest dose positive in any one of the strains involved in one entry is taken.

test batteries for detecting mutagenic carcinogens. Parodi et al. (*7*) have proposed a method using several parameters to predict both qualitatively and quantitatively the carcinogenic activity of chemicals. The success of this approach was found to be chemical-class dependent.

The committee 1 activity to date has been directed toward ranking for mutagenic activity. Future efforts are planned for comparing the ICPEMC mutagenicity rankings to animal carcinogen standards such as those proposed by Gold et al. (*8*). In an activity related to this end, Nesnow (*9*) constructed a multifactor ranking scheme for comparing the carcinogenic activity of chemicals. This scheme was produced in collaboration with committee 1 and used a similar process to weight factors that influence potency to the one used in the mutagenicity ranking approach.

Each of the methods described has attributes that make it useful for specific purposes, but the methods are all primarily

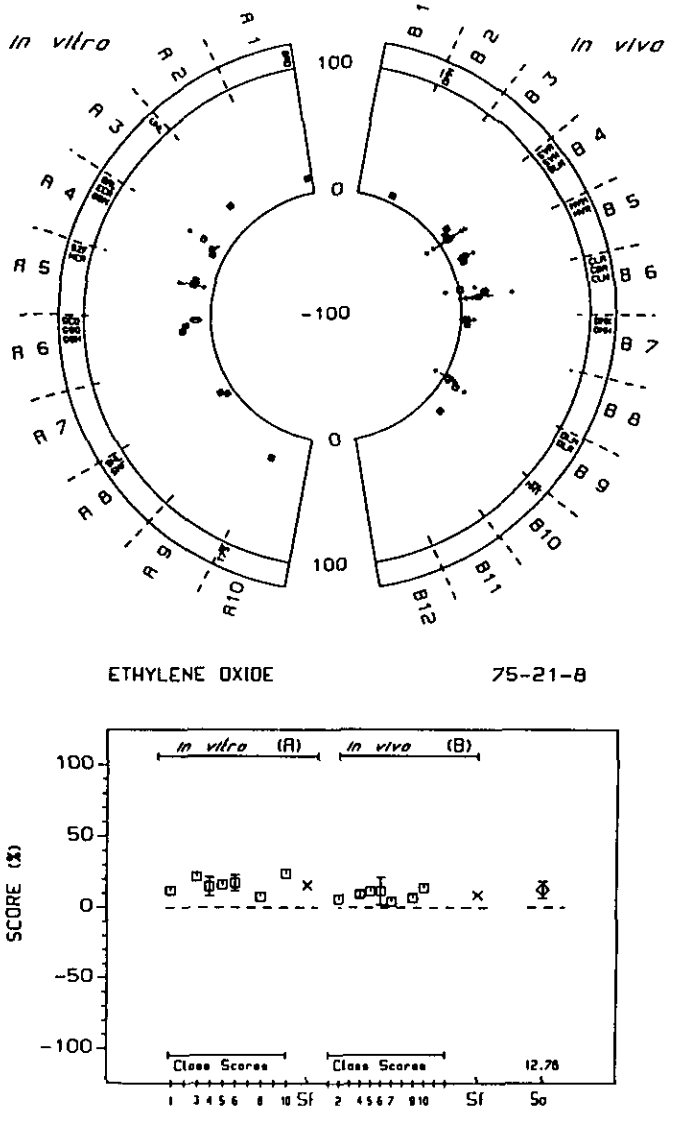


FIGURE 2. The ICPEMC mutagenic activity profile for ethylene oxide. The upper portion of the diagram gives the individual test results for *in vitro* and *in vivo* assays. The location of the response on the scale (−100 to +100) indicates whether the results were positive or negative. The lower portion of the diagram gives the class (1–10), family (Sf), and agent (Sa) scores. Classes are identified by the numbers along the *X*-axis.

oriented toward carcinogenesis. GAPs are similar to the committee 1 approach in both graphic output and in the fact that they are directed toward mutagenicity *per se*.

Data Evaluation Methods Developed by ICPEMC

Once the basic structure of the committee 1 approach had been determined, data collection and analysis programs were written in FORTRAN 77 for a Digital VAX 750 computer. The software program was designed to be flexible and amenable to adjustment (fine tuning) as data entered into the database were evaluated. An alternate version of the program is being prepared for IBM-AT compatible personal computers. The ICPEMC approach has been identified as the mutagenic activity profile (MAP) method because of the graphic output format and because the scheme ranks chemicals according to their activity. Details of the data evaluation system and the techniques employed to maximize use of the method are currently in press (*10,11*).

In summary, the approach uses a weight-of-evidence concept combined with unweighted averaging of modified test results. The qualitative test responses (positive or negative) are modified by two factors: dose and assay replication. Defining doses are selected from the lowest effective dose (positive results) and the highest ineffective dose (negative results). Dose modifiers, which have been corrected for bias introduced by characteristics associated with the test system (*11*), are then applied to the calculations.

Each test system for which data can be entered into the scheme is uniquely identified by a three-letter code (Table 1) proposed by Waters et al. (*4*). Trials of individual tests are transformed to produce test scores. Scores from individual tests are combined into class scores by simple unweighted averaging. Test classes have phylogenetic and end point traits in common (e.g., gene mutation tests in prokaryotic cells, chromosome aberrations in cultured mammalian cells); a class such as A6 consists of tests that are presumed to detect gene mutation in cultured mammalian cells. Results from the L5178Y mouse lymphoma assay, HGPRT assay in Chinese hamster ovary or V79 cells, or gene mutation tests using human cell types would be combined in the A6 class. *In vivo* classes were constructed in a similar fashion. For example, class B6 consists of bone marrow metaphase cytogenetic analysis in mice, rats, hamsters, and humans.

Merging data into classes is performed by simple averaging. Class scores are combined into family scores, again by simple averaging. There are two family scores, one for *in vitro* results and one for *in vivo* results.

Figure 1 summarizes the steps in the process for assembling and merging data into test, class, family, and agent scores. The process determines a score for each trial of a given test and then merges them into a score for the test, a score for the class, a score for the family, and finally, a single agent score (Sa) representing the consensus (weight-of-evidence) for the chemical. The consensus score defines the overall mutagenic activity based on all the test results.

The results of the evaluation process are expressed in both tabular and graphic formats. The tabular output lists each of thescores identified above, the calculations producing the scores, and reference citations for each of the data entries. The graphic format for ethylene oxide (Fig. 2) is used as an example and can be compared to the GAP graphics in Figure 3. The ICPEMC profiles are presented in diagrams with upper and lower plots. The upper portion of the diagram gives (in the two hemispheres) modified test scores for each trial (with a mean and confidence limits if the replicate number is three or greater), along with the three-letter identification code. Agent scores (Sa) can theoretically range from −100 to +100 with the 0 separating the active (+) or inactive (−) responses.

At each step of the process, scores are averaged with negative results down-weighting positive scores. The major determinants for location of the scores on scale are sign (+ or −), defining dose, and replication of the test. The final merging represents a consensus of all entries. The test codes are arranged so that the

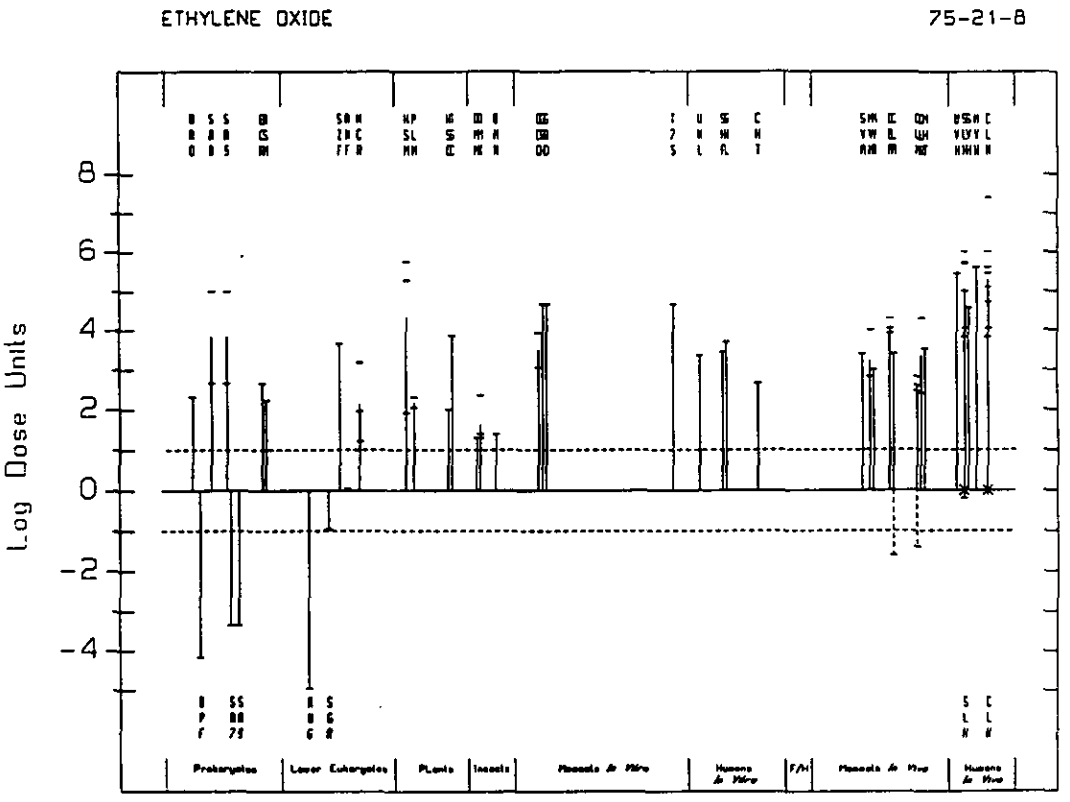


FIGURE 3. The genetic activity profile for ethylene oxide. Tests are arranged in phylogenetic groups. Data are formatted according to the strength of response in both positive and negative directions. (Courtesy of M. Waters.)

tests within a given class (e.g., A1, A2, or B1, B2) are clustered together. The lower portion of the diagram provides the class scores, family scores, and agent score. The name of the chemical, current date, and CAS number for the agent are also provided on the plot.

The rationale for including a graphic as well as tabular outputs are *a*) to provide all data in a convenient, informative manner on a single page for quick reference and *b*) to permit users to follow the influence of the data reduction steps on the initial test results.

The data analysis and merging program has continued to evolve as more insight about test performance and data analysis has been gained. Consequently, there have been several versions of the agent scores, which have resulted in slight shifts of the chemical ranking. The system is approaching a point where the committee believes that it is working sufficiently well that final settings for the modifiers can made and the system should be released for general use. Because of the design of the program, additional information gained during use of the system can be used to "educate" the process by fine tuning the modifiers or by weighting some of the variables (*10,12*).

In developing the process in this manner, certain assumptions were made by the members of the committee: *a*) there were no established procedures available for using test results to classify chemicals as nonmutagens, but one was needed; *b*) there was insufficient information available to set weights for different tests. Therefore, all tests were assumed to be equally relevant to the process of determining mutagenic activity; *c*) both *in vitro* and *in vivo* data would be required to provide an accurate assessment of the genetic activity of a chemical; *d*) replication of the agent in a test (up to a point) should provide, on the average, a better estimate of the mutagenic activity for the chemical than a single trial; *e*) merging test results, especially replicates of a test and tests measuring the same end point in similar types of organisms, would not significantly violate scientific principles because a similar process is performed intuitively by most toxicologists when evaluating multitest results for a chemical.

Source of Data in the Database

The current database used to evaluate the approach and perform the statistical analyses consists of 4490 results for 113 chemicals. The primary data was provided to ICPEMC by the U.S. Environmental Protection Agency and contained results from many of the chemicals in the IARC Supplement 6 (*13*). The chemicals in the MAP database all have at least three *in vitro* tests and at least two *in vivo* tests. The committee set these minimums as requirements to evaluate the ability of the method to handle large heterogeneous data sets and because most of the test batteries in common use generally contained both *in vitro* and *in vivo* tests.

Concerns and Limitations of the Approach

The committee realized that developing a data evaluation scheme would involve treating genotoxicity data in ways that are different from treatments typically used to evaluate groups

Table 2. Agent scores for 113 chemicals in the database.

Chemical	Score	Chemical	Score
Ethanol	−27.70	Vinyl chloride	0.20
Melamine	−26.38	Acrylonitrile	0.54
Chlorodifluoromethane	−26.05	*p*-Nitro-*o*-phenylenediamine	0.58
C.I. Acid Red 14	−23.26	Diethylstilbestrol	0.76
Pentachloronitrobenzene	−20.45	Malonaldehyde	0.79
Saccharin	−18.78	2,3,7,8-TCDD	1.38
Halothane	−18.69	1-Naphthylamine	1.46
Inoniazide	−18.59	Vinylidene chloride	1.74
Phenylbutazone	−18.39	Auramine	2.11
Caprolactam	−18.21	Cadmium	2.44
Diethylhexylphthalate	−17.35	Methotrexate	2.73
Polychlorinated biphenyls	−16.88	2,4-D	3.14
Ethylenethiourea	−16.77	MCPA	3.29
Sodium saccharin	−15.76	Aldrin	3.34
Methoxychlor	−15.71	Procarbazine HCl	3.36
Polybrominated biphenyls	−15.46	Benzyl chloride	3.58
Chloroform	−15.28	Dimethlycarbamoyl chloride	4.87
Chloramphenicol	−15.11	Azathioprine	4.89
Metronidazole	−14.75	Dibromochloropropane	5.71
Maleic hydrazide	−14.13	Nickel	5.78
1,1,1-Trichloroethane	−14.05	Benzidine	5.88
Dichloromethane	−13.85	Hycanthone methanesulfonate	5.95
Tetrachloroethylene	−13.48	Acetaldehyde	6.05
Phenobarbital	−13.22	Ethylene dibromide	6.60
Endrin	−12.81	Diethyl sulfate	7.11
Mestranol	−12.43	*Tris*(2,3-dibromopropyl)PO_4	7.77
Progesterone	−12.11	Propylene oxide	7.80
Tetraethylthiurum disulfide	−11.92	Arsenic(III)	8.04
Malathion	−11.32	Hydrazine	8.30
Amitrole	−10.68	Styrene oxide	8.49
N-Nitrosodiphenylamine	−9.74	2-Naphthylamine	9.11
Aniline	−8.89	Benzo(*a*)pyrene	9.52
Lead	−8.72	Formaldehyde	9.74
Chrysene	−8.58	Myleran	9.96
Asbestos	−8.39	Vincristine sulfate	10.42
Benzene	−7.15	Epichlorohydrin	10.60
Caffeine	−6.67	Uracil mustard	11.09
Sodium fluoride	−5.90	Cyclophosphamide	11.30
Cyclohexylamine	−5.83	6-Mercaptopurine	12.32
Heptachlor	−5.06	Ethylene oxide	12.78
Diazepam	−4.68	Dimethyl sulfide	13.92
DDT	−4.30	1-(2-Chloroethyl)-3-cyclohexyl-1-nitrosourea	14.17
Carbon tetrachloride	−4.17	Chlorambucil	14.55
Methyl parathion	−3.85	Bleomycin	16.90
Dieldrin	−3.67	Vinblastine sulfate	18.19
Phenytoin	−3.31	Chloroprene	18.32
o-Toluidine	−3.29	MNNG	18.32
Trichloroethylene	−2.34	Methyl bromide	18.33
Benz(*a*)anthracene	−1.45	Chromium(IV)	19.01
Styrene	−1.28	BCNU	19.48
Hexachlorocyclohexane	−1.15	8-Methoxypsoralen (+UVR)	19.81
Pentachlorophenol	−0.51	Melphalan	23.07
Dimethoate	−0.31	Actinomycin *D*	23.10
5-Fluorouracil	−0.26	*Cis*platen	23.31
		Aflatoxin B_1	24.67
		Thiotepa	25.91
		Nitrogen mustard	26.70
		Adriamycin	29.22
		Triaziquone	49.67

of test results. For example, the process of averaging test and class scores was seriously questions because of the concern that a single, possibly highly relevant, test result would be diluted by larger numbers of negative results. This potential problem was emphasized because of another limitation expressed and that was that input of data does not require prior expert review, thus a positive result from a well-performed test may be masked by several studies not properly performed with negative results. There was less concern that the converse of this situation might occur.

Another concern expressed by committee members as well as commission members reviewing the approach was the decision to give equal weight to *in vitro* and *in vivo* tests. *In vivo* data are generally viewed as more relevant to hazard identification and typically given more weight.

Many individuals reviewing the process questioned the rationale for merging data by simple averaging of modified scores. This not only raised the potential of diluting unique test responses as indicated earlier but was also of concern because there was a general belief that tests measuring different genetic end points (gene mutation, aberrations, sister chromatid exchange, transformation, etc.) measure quite different mechanistic phenomena that cannot be merged by simple averaging. There were other concerns of a lesser nature that were identified and recognized by the committee during its deliberations over the past several years.

The committee members considered all of these concerns and other likely limitations during the construction of the MAP scoring system. Resolution of all questions was not possible, but the output of the scoring system with the existing data suggested in several cases that the potential limitations did not seriously flaw the evaluation scheme.

Results

Even with the limitations encountered, the MAP system produced by ICPEMC appears to accomplish many of the goals initially stated by the committee. Table 2 is a listing of the rank order 113 chemicals used in constructing the database. Some additional fine tuning of the system is expected, and before final release there could be some minor changes in the rank order of agents. In this latest version, ethanol, with an agent score of −27.70 (Fig. 4), was the least genetically active agent in the database, and triaziquone (Trenimon) with an agent score of +49.67 (Fig. 5), was the most genetically active. The rank order, with a few exceptions, seems consistent with an intuitive ranking of mutagenic activity or with rankings from other experts or expert systems.

The number of test entries per chemical ranged from a low of

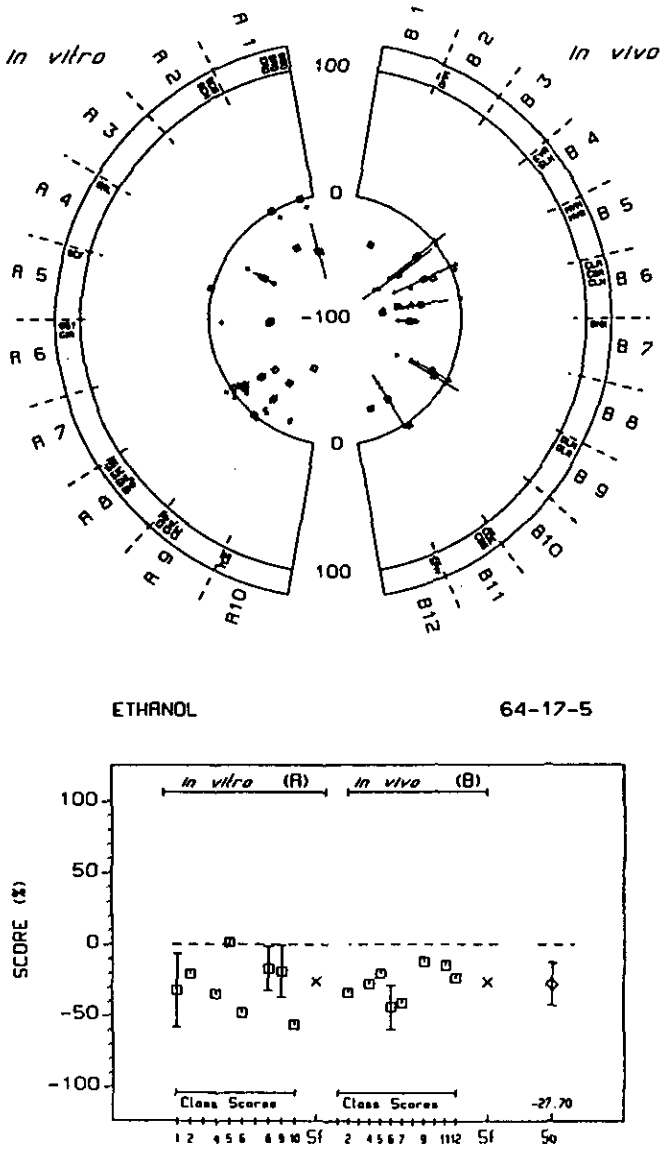


FIGURE 4. The ICPEMC mutagenic activity profile for ethanol.

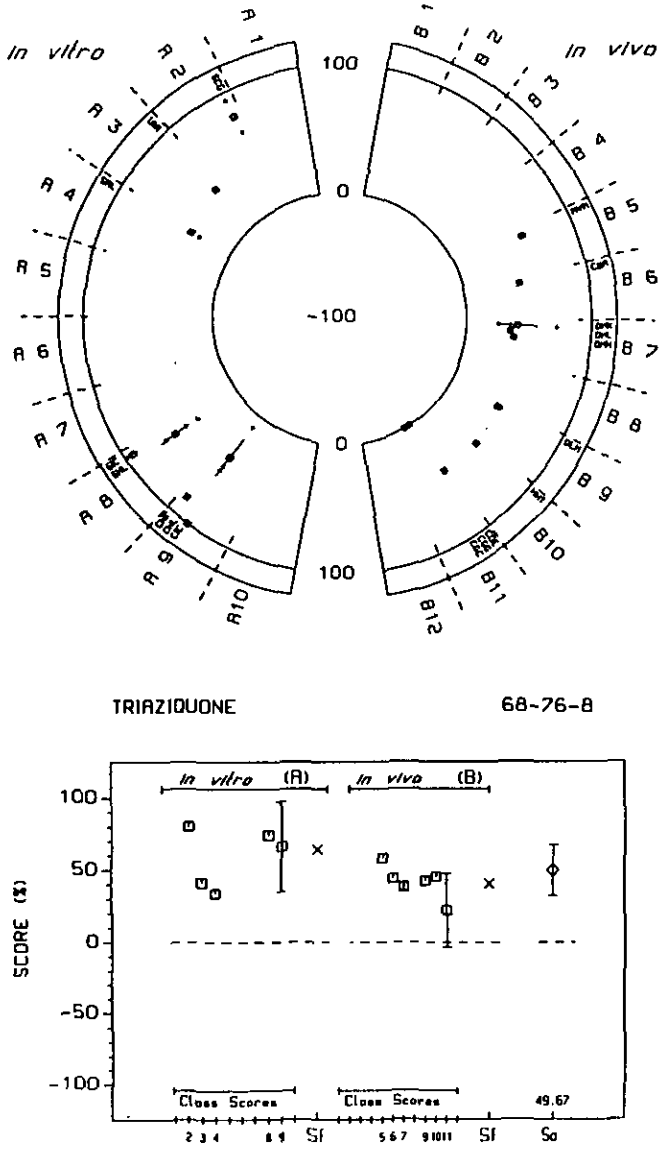


FIGURE 5. The ICPEMC mutagenic activity profile for triaziquone.

6 studies for mestranol to 275 studies for cyclophosphamide. Among the 113 data sets, 108 (96%) had mixed test results (both positive and negative). From the data available at the time of this report, only C.I. acid red (11 entries), melamine (8 entries), mestranol (6 entries), and polybrominated biphenyls (13 entries) consisted of entirely negative test data. Only chloroethyl-cyclohexyl-nitrosourea (9 entries) had all positive test results.

Data Interpretation

To fully use the MAP system, a practical application of agent (Sa) scores must be developed. One can define, on a limited basis, the activity of a chemical (e.g., mutagen, clastogen) from the unequivocal, reproducible data from a single test system such as the Ames test, the Drosophila sex-linked recessive lethal assay, or the mouse micronucleus assay; however, such a definition

Table 3. Agent scores for chemicals reported to be carcinogenic to humans.

Chemical	Agent score
Asbestos	−8.39
Benzene	−7.15
Vinyl chloride	+ 0.20
Diethylstilbestrol	+ 0.76
Azathioprine	+ 4.89
Benzidine	+ 5.88
Arsenic	+ 8.04
2-Naphthylamine	+ 9.11
Myleran	+ 9.96
Cyclophosphamide	+11.30
Chlorambucil	+14.55
Chromium(VI)	+19.01
8-Methoxypsoralen (+UVR)	+19.81
Melphalan	+23.07
Nitrogen mustard	+26.70

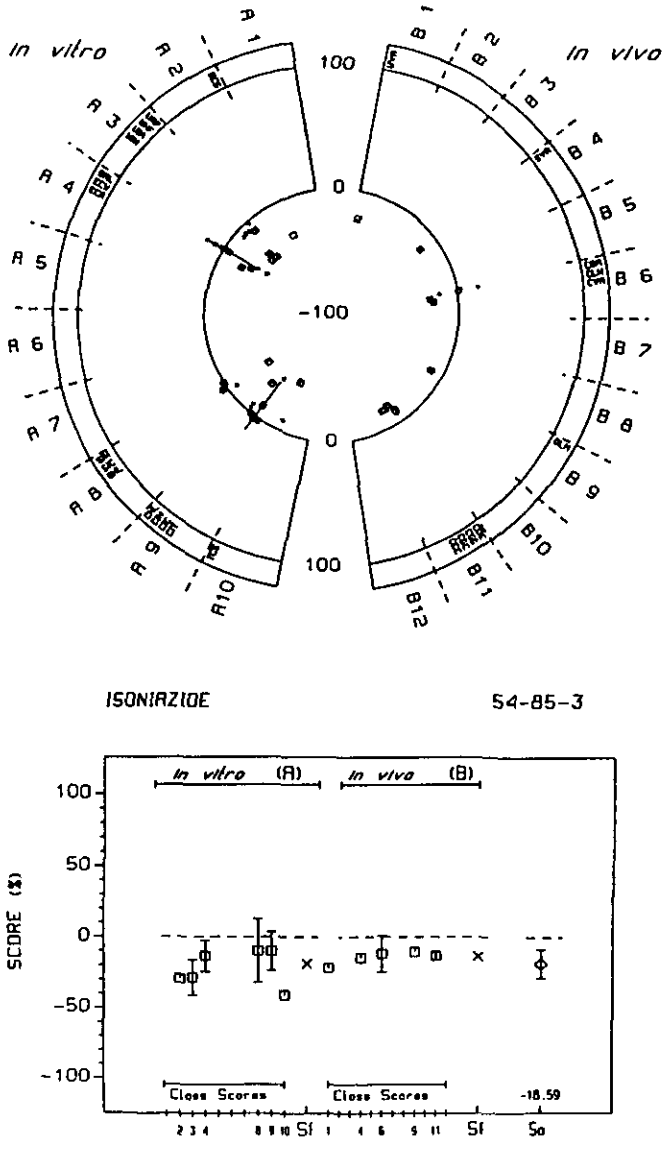


FIGURE 6. The ICPEMC mutagenic activity profile for isonaizide.

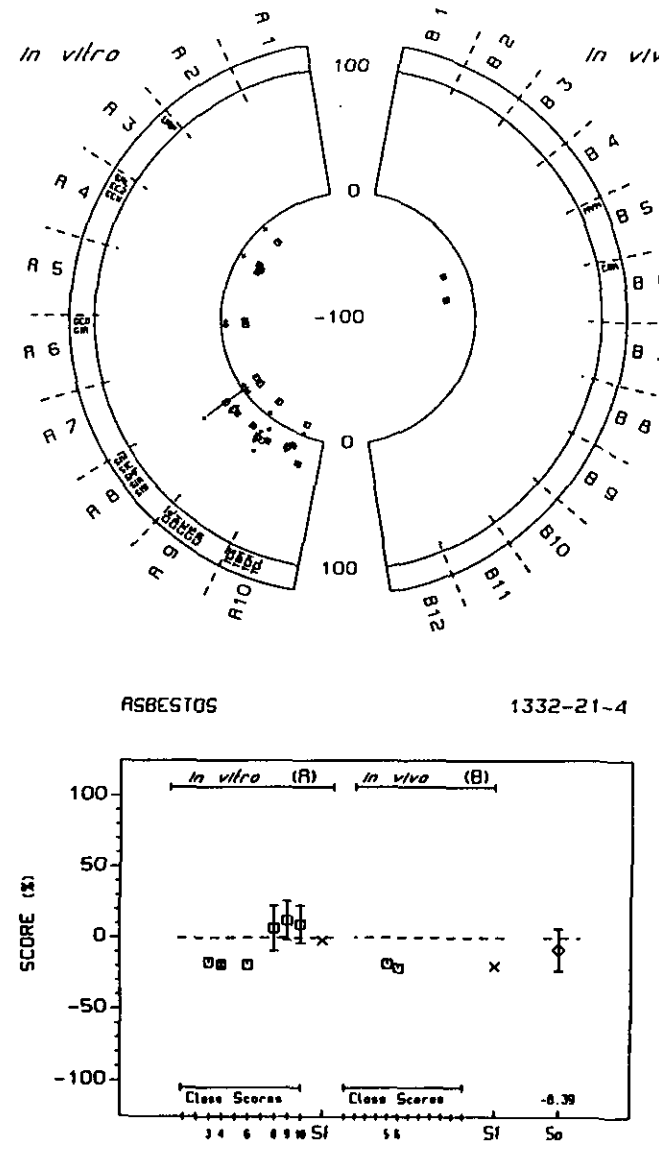
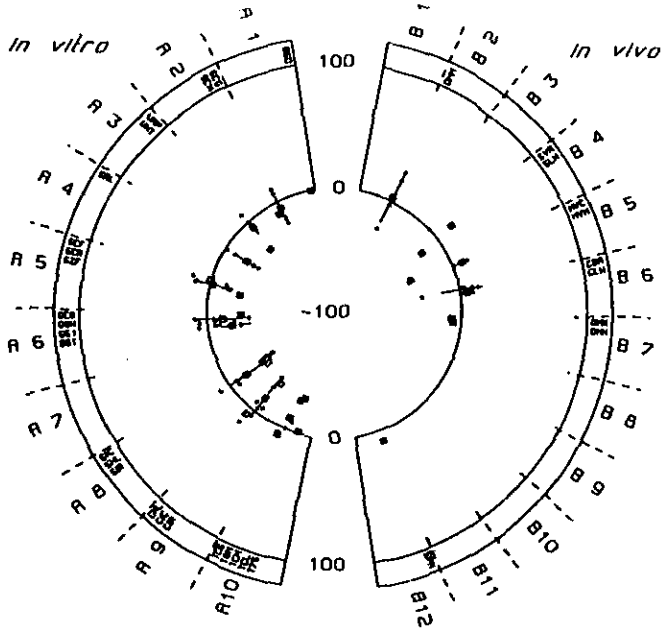


FIGURE 7. The ICPEMC mutagenic activity profile for asbestos.

carries little information concerning how generalized the activity might be across other test methods or other species and has no quantitative indication of potency. The ICPEMC scoring system attempts to introduce these two attributes into the mutagenicity definition. Several uses for the agent score have been considered as discussed below.

The agent score could be viewed as an indication of the level of confidence (probability) that a chemical is a "general" mutagen across test and species boundaries. In other words, how likely is the chemical to produce a positive or negative response in the next assay to which it is subjected? The higher the agent score, the greater the probability that the chemical is a "general" mutagen and represents a human hazard. Agents that show potent but highly test-method-specific responses (i.e., a single test positive) will not generate a high agent score. Consequently, the agent score from a test battery could serve as a quantitative estimate of the genetic hazard of a compound.



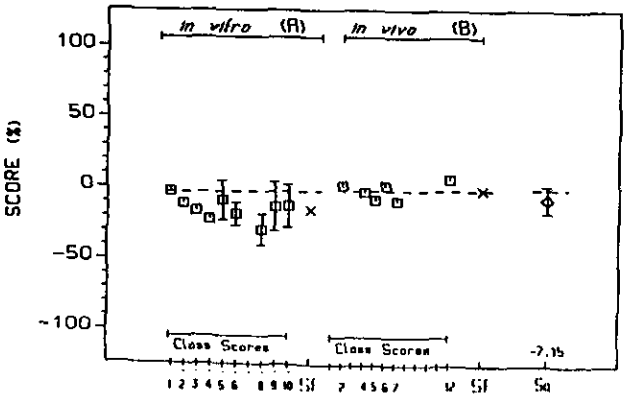
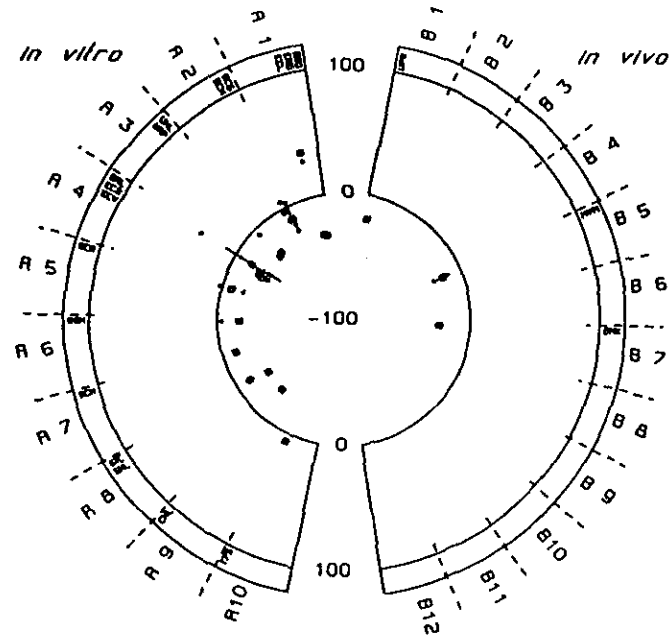


FIGURE 8. The ICPEMC mutagenic activity profile for benzene.



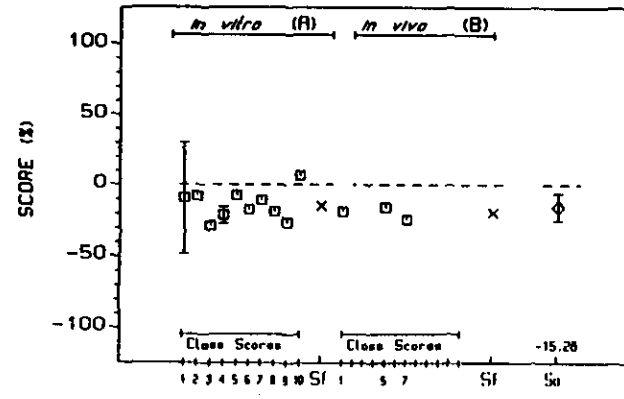


FIGURE 9. The ICPEMC mutagenic activity profile for chloroform.

The agent score might be used in a qualitative manner to establish potential for germ cell hazard. Among the 113 chemicals in the database, 8 have been reported positive in rodent tests for heritable germ cell effects (*14,15*). Seven of the 8 (88%) germ cell mutagens showed positive agent scores. The one compound designated a germ cell mutagen which had a negative agent score was isoniazid (Fig. 6). A weak positive effect was reported in the mouse heritable translocation assay (*1*).

Some consideration has also been given to the use of the agent score as an indicator of carcinogenic potential. Fifteen of the 113 chemicals fall into the IARC group I human carcinogens (*16*). Thirteen of the 15 (87%) have positive agent scores (Table 3).

The two human carcinogens with negative agent scores are asbestos (Fig. 7) and benzene (Fig. 8). Attempts to use the agents score rankings to predict rodent carcinogenesis potency have resulted in several conflicts with conventional judgments. Al-

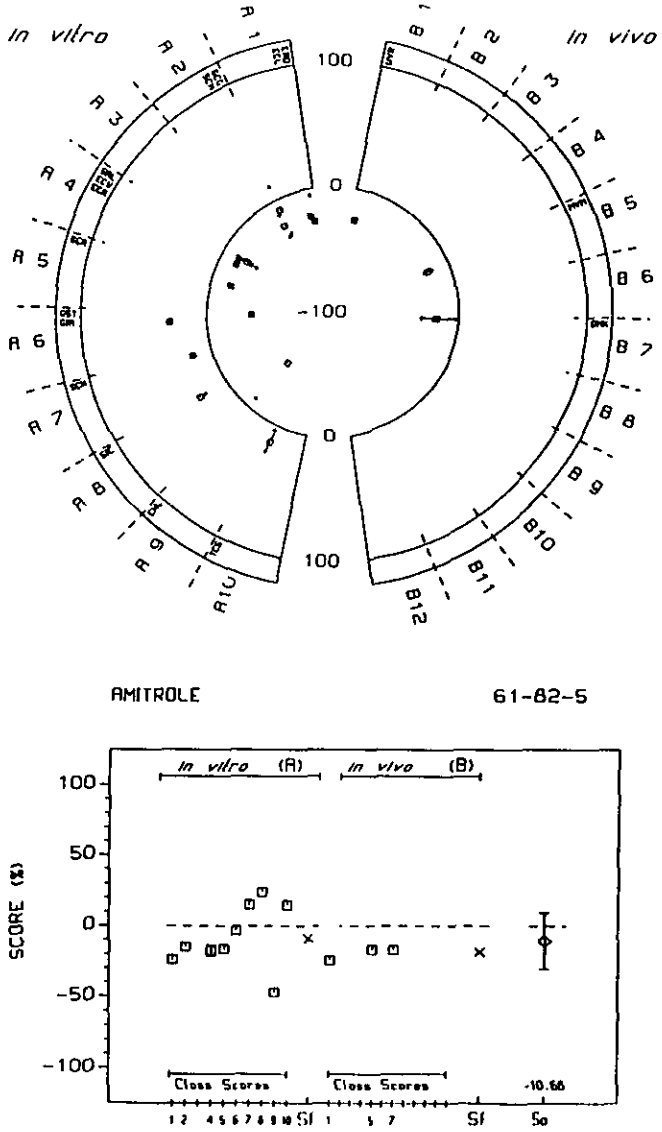


FIGURE 10. The ICPEMC mutagenic activity profile for amitrole.

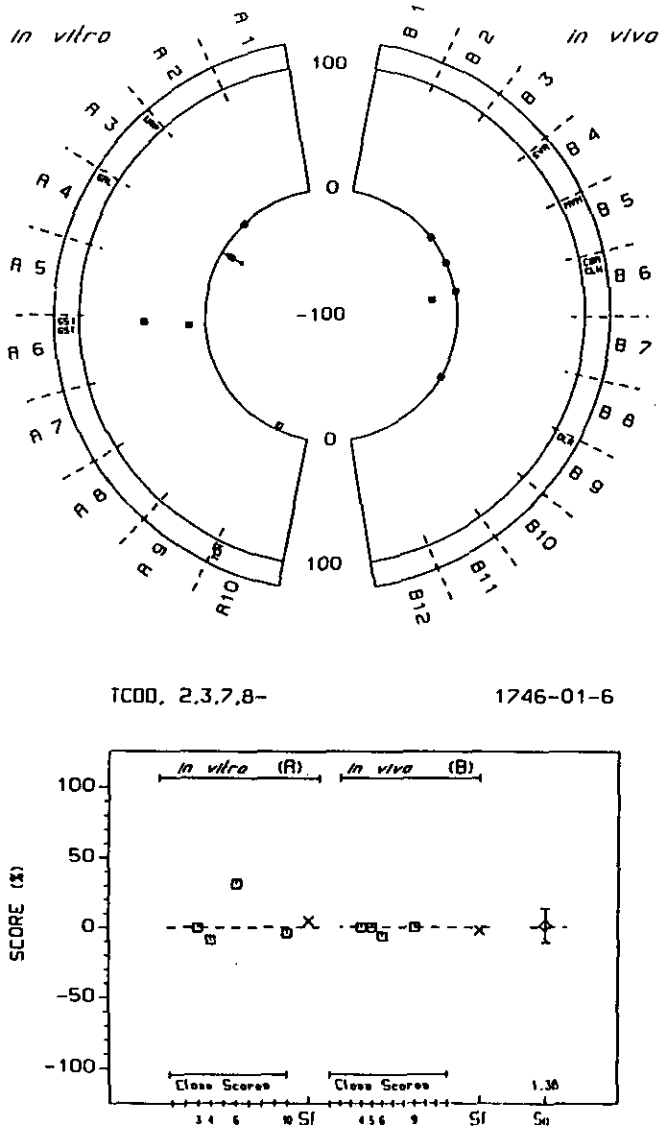


FIGURE 11. The ICPEMC mutagenic activity profile for TCDD.

though many rodent carcinogens fall among the chemicals with high agent scores, some highly active rodent carcinogens such as chloroform (Fig. 9), amitrole (Fig. 10), and TCDD (Fig. 11) all exhibited low agent scores. These agents belong to a heterogenous group of chemicals whose mechanisms of carcinogenesis are believed to be other than genotoxic (*17*). A subset of the 113 chemicals with these characteristics is listed in Table 4. Seventeen of the 19 agents in this nongenotoxic category have negative agent scores consistent with their assumed mechanisms and are also not mutagenic in the conventional Ames assay.

The committee is currently evaluating the alternative uses of the agent scores. The relative ranking of chemicals in Table 2 coincides reasonably well with an intuitive assessment of their genetic hazard. This is especially true for those with very high or very low agent scores. There appear to be a few anomalies among the chemicals in the database, for example, procarbazine.

Table 4. Agent scores for chemicals considered to produce tumors in rodents by nongenotoxic mechanisms.

Chemical	Agent score
Diethylhexylphthalate	−17.35
PCBs	−16.88
Ethylenethiourea	−16.77
PBBs	−15.46
Chloroform	−15.28
1,1,1-Trichloroethane	−14.05
Tetrachloroethylene	−13.48
Phenobarbital	−13.22
Endrin	−12.81
Progesterone	−12.11
Amitrole	−10.68
Asbestos	−8.39
Heptachlor	−5.06
DDT	−4.30
Carbon tetrachloride	−4.17
Dieldrin	−3.67
Trichloroethylene	−2.34
Diethylstilbestrol	+0.76
TCDD	+1.38

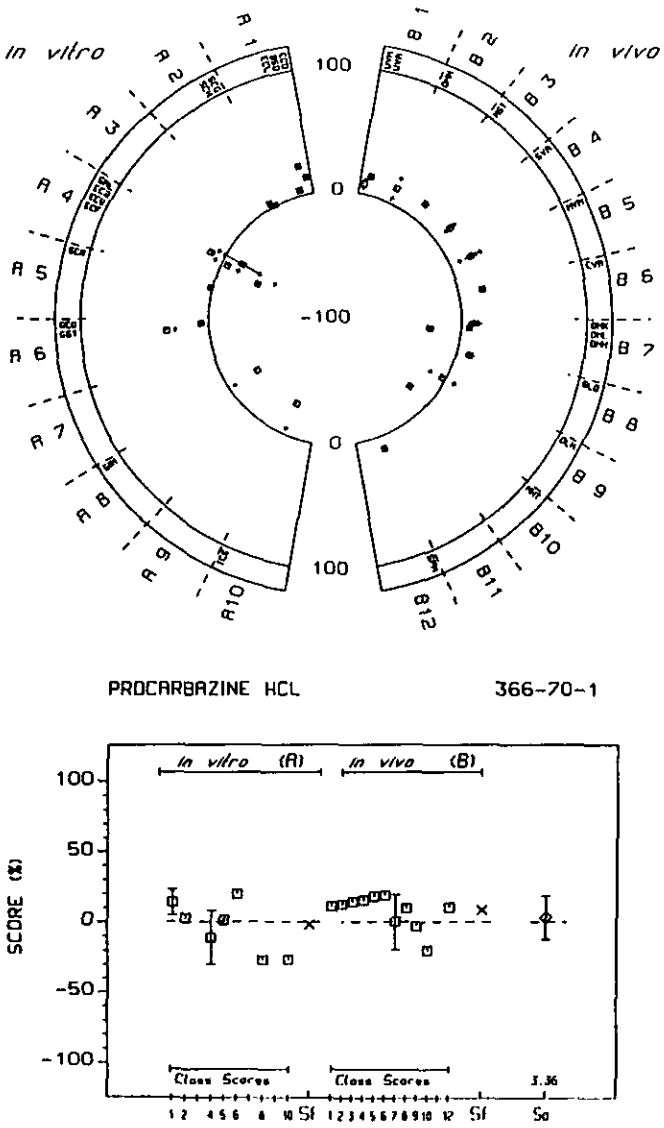


FIGURE 12. The ICPEMC mutagenic activity profile for procarbazine.

Hydrochloric acid (Fig. 12) has a relatively low agent score of 3.36. This chemical is highly mutagenic in rodent germ cells (*18*), yet ranks lower than other agents that would presumably pose less of a genetic risk (e.g., acetaldehyde, nickel and formaldehyde). Benzene, which is quite active as a clastogen *in vivo*, has an agent score of −7.15. This anomaly appears to result from the fact that a large number of negative studies have been conducted *in vitro* and these have diluted the limited number of positive results *in vivo*. This is an example related to some of the concerns expressed earlier. Both procarbazine and benzene appear lower in the agent score rankings than might be presumed generally. Few instances of this situation were found upon an extensive analysis of the database.

Conclusions

In spite of the early stage of development, it is clear that the ICPEMC committee 1 MAP approach of integrating and processing genetic toxicology data is capable of meeting many of the initial requirements set forth by the committee. The approach is able to cope with redundant, disparate, and missing data in the published literature.

From the current database of 113 chemicals, the scoring method in its current configuration was capable of correctly assigning scores to almost all of the known heritable mutagens. Most human carcinogens in the database were assigned positive agent scores, and the category of rodent carcinogens presumed to induce tumors by nongenotoxic mechanisms were all assigned negative agent scores by the method.

A crucial element in this exercise was to compare the mutagenic ranking of chemicals with their ranking as rodent carcinogens. To accomplish this, a parallel system for rank-ordering rodent carcinogens was developed by Nesnow (*19*). Once this new database is filled with sufficient chemicals to make a comparison meaningful, the results will be published.

A comprehensive statistical analysis has been performed with the existing database (*11*). Several preliminary findings have produced important insight into mutagenicity testing: *a*) *In vitro* and *in vivo* tests appear to respond similarly to a broad range of chemicals. *b*) Chemicals do not appear to be highly specific for genetic end points (gene mutation, sister chromatid exchange, clastogenicity, cell transformation). Class scores proved to be very congruent with the consensus (Sa) scores for the 113 chemicals. *c*) Using the 113 chemicals as surrogates for the universe of chemicals, the range of agent scores fall generally on a continuous, rather than a bimodal, scale with approximately half the chemicals having positive agent scores and half having negative agent scores.

The study and refinement of the ICPEMC committee 1 MAP method of complex mutagenicity data evaluation will continue. Its adaptation to data assessment will be enhanced by the availability of software modified for use on personal computers. Based on the initial experiences with the approach, it is clear that important insights about genetic tests and test batteries will emerge. Whether this approach will break through the current barriers encountered in using genetic test to predict carcinogenicity remains to be seen.

The author recognizes ICPEMC for its support of the Committee's work on this project. Other ICPEMC committee 1 members are Paul Lohman, Mort Mendelsohn, Mike Waters, and John Ashby. Additional statistical and programming assistance has been provided by Dan Moore, II and Walter Lohman.

REFERENCES

1. ICPEMC, Committee 1 Final Report. Screening strategy for chemicals that are potential germ-cell mutagens in mammals. Mutat. Res. 114: 117–177 (1983).
2. Brusick, D. Unified scoring system and activity definitions for results from in vitro and sub-mammalian mutagenesis test batteries. In: Health Risk Analysis (E. D. Copanhaven, C. R. Richmond, and P. J. Walsh, Ed.), Franklin Institute Press, Philadelphia, 1981, pp. 273–286.
3. Squire, R. A. Ranking animal carcinogens: a proposed regulatory approach. Science 214: 877–880 (1981).
4. Waters, M. D., Stack, H. F., Brady, A. L., Lohman, P. H. M., Vanio, H. Use of computerized data listings and activity profiles of genetic and related effects in the review of 195 compounds. Mutat. Res. 205: 295–312 (1988).
5. Benigni, R., and Giuliani, A. Predicting carcinogenicity with short-term tests: biological models and operational approaches. Mutat. Res. 205: 227–236 (1988).
6. Chankong, V., Haimes, Y. Y., Rosenkranz, H. S., and Pet-Edwards, J. The carcinogenicity prediction and battery selection (CPBS) method: a Bayesian approach. Mutat. Res. 153: 135–166 (1985).

7. Parodi, S., Taningher, M., Romano, P., Grilli, S., and Sanit, L. Mutagenic and carcinogenic potency indices and their correlation. Teratog. Carcinog. Mutagen. 10: 177–197 (1990).
8. Gold, L. S., Sawyer, C. B., Magaw, R., Backman, G. M., de Veciana, M., Levinson, R., Hooper, N. K., Havender, W. R., Berstein, L., Peto, R., Pike, M. C., and Ames, B. N. A Carcinogenic Potency Database of the standardized results of animal bioassays. Environ. Health Perspect. 58: 9–319 (1984).
9. Nesnow, S. A multi-factor carcinogen potency ranking scale for comparing the activity of chemicals. Mutat. Res. 239: 83–115 (1990).
10. Lohman, P. H. M., Mendelsohn, M. L., Moor, D. H., II, Waters, M. D., Brusick, D. J., Ashby, J., and Lohman, W. A method for comparing, combining and interpreting short-term genotoxicity data: the basic system. Mutat. Res., submitted.
11. Moore, D. H., II, Mendelsohn, M. L., and Lohman, P. H. M. A method for comparing, combining and interpreting short-term genotoxicity data: the optimal use of dose information. Mutat. Res., submitted.
12. Mendelsohn, M. L., Moore, D. H., II, and Lohman, P. H. M. A method for comparing, combining and interpreting short-term genotoxicity data: results and interpretation. Mutat. Res., submitted.
13. IARC. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Supplement 6, Genetic and Related Effects: An Updating of Selected IARC Monographs from Volumes 1–42. International Agency for Research on Cancer, Lyon, 1987.
14. Russell, L. B., Selby, P. B., Von Halle, E., Sheridan, W., and Valcovic, L. The mouse specific-locus test with agents other than radiations. Interpretation of data and recommendations for future work. Mutat. Res. 86: 329–354 (1981).
15. Ehling, U. H. Quantitation of the genetic risk of environmental chemicals. Risk Anal. 8: 45–57 (1988).
16. IARC. Chemicals and Industrial Processes Associated with Cancer in Humans, Supplement 1, Volumes 1–20. International Agency for Research on Cancer, Lyon, 1979.
17. Williams, G. M., and Weisberger, J. H. Chemical carcinogens. In: The Basic Science of Poisons, 3rd ed. (C. Klaassen, Ed.), Macmillan, Inc., New York, 1986, pp. 99–173.
18. Ehling, U. H., and Neuhauser, A. Procarbazine-induced specific-locus mutations in mice. Mutat. Res. 59: 245–256 (1979).